

On the interpolation of calibration solutions obtained in radio interferometry

Sarod Yatawatta^{1*}

¹*ASTRON, The Netherlands*

26 September 2012

ABSTRACT

Full polarimetric radio interferometric calibration is performed by estimating 2 by 2 Jones matrices representing instrumental and propagation effects. The solutions obtained in this way differ from the true solutions by a 2 by 2 unitary matrix ambiguity. This ambiguity is common to all stations for which a solution is obtained but it is different for solutions obtained at different time and frequency intervals. Therefore, straightforward interpolation of solutions obtained at different time and frequency intervals is not possible. In this paper, we propose to use the theory of quotient manifolds for obtaining correct interpolants that are immune to unitary matrix ambiguities.

Key words: Instrumentation: interferometers; Methods: numerical; Techniques: interferometric

1 INTRODUCTION

Most modern radio interferometers have dual polarized feeds and therefore, the use of the matrix measurement equation (Hamaker et al. 1996) gives a compact and accurate description of their operation. Calibration of such an interferometer essentially boils down to the estimation of Jones matrices of size 2×2 with complex entries. As shown by Hamaker (2000), the solutions acquired for the Jones matrices will only be equivalent to the true solution upto a unitary matrix ambiguity.

This ambiguity would not hinder further processing of data because it cancels out during correction of the data using the obtained solutions. However, the ambiguities do prevent us from using the solutions for further modeling of instrumental effects (e.g., beam shape (Yatawatta 2012)) and effects due to the propagation medium such as the ionosphere (Intema et al. 2009). Moreover, interferometers operating at low frequencies have a wide field of view and calibration has to be performed along hundreds of directions in an efficient manner (Kazemi et al. 2011). Along each direction, we would have an ambiguity which is independent of other directions.

In this paper, we present a method of interpolation of the calibration solutions (Jones matrices) and we consider the case where each solution is affected by any unknown unitary ambiguity. We consider the simplest case of interpolation to present our method: Finding the mean of a given set of solutions which can also be extended to interpolation with weighted averaging. The uses of interpolation or aver-

aging are numerous. First, averaging of solutions obtained at adjacent time and frequency intervals provides us with a robust estimate of calibration solutions especially under noisy situations. Moreover, interpolation reduces the number of data points that needs to be visualized. This is important when we have solutions over hundreds of directions in the sky at a number of time and frequency intervals. Interpolation also provides us solutions when there is missing or flagged data points.

Traditional calibration software such as AIPS that are based on a scalar data model has the ability to interpolate scalars such as the amplitude or phase of a single polarization. However, this is not possible when we are dealing with 2×2 matrices as what is presented in this paper. Due to the unitary ambiguity, any linear operation in Euclidean space with such matrices would not give us a feasible interpolant. Therefore, we explore the quotient manifold structure (Absil et al. 2008) of the calibration solutions. We use the theory of interpolation over manifolds to get a feasible solution to our problem. We present an algorithm to find the mean of a given set of calibration solutions. This algorithm could also be extended to interpolation with any positive weighting scheme. The interpolant we obtain using this algorithm still has a unitary ambiguity but it is closer to the true interpolant (without ambiguities) than what we obtain by standard (Euclidean) interpolation. Similar work on averaging or interpolation over manifolds and Lie groups appear in diverse areas of research and we refer the reader to Fiori & Tanaka (2009), Fiori (2011), Kaneko et al. (2012), and Amsellem & Farhat (2008) for further information.

The rest of the paper is organized as follows: In section 2, we give an overview of radio interferometric calibration

* E-mail: yatawatta@astron.nl

and the ensuing ambiguities. In section 3 we present the interpolation scheme. We provide simulation results to prove the superiority of the proposed scheme in section 4 and finally, draw our conclusions in section 5.

Notation: Matrices and vectors are denoted by bold upper and lower case letters as \mathbf{J} and \mathbf{v} , respectively. The transpose and the Hermitian transpose are given by $(\cdot)^T$ and $(\cdot)^H$, respectively. The matrix Frobenius norm is given by $\|\cdot\|$. The set of complex numbers is denoted by \mathbb{C} . The identity matrix is given by \mathbf{I} . The angle of a complex number is given by \angle .

2 RADIO INTERFEROMETRIC CALIBRATION

Consider a radio interferometer with N stations. Let the sky signal consist of M discrete sources. The observed data on the baseline pq formed by stations p and q is given by (Hamaker et al. 1996)

$$\mathbf{V}_{pq} = \sum_{m=1}^M \mathbf{J}_{pm} \mathbf{C}_{pqm} \mathbf{J}_{qm}^H + \mathbf{N}_{pq} \quad (1)$$

where \mathbf{V}_{pq} and \mathbf{N}_{pq} ($\in \mathbb{C}^{2 \times 2}$) are the visibility matrix, and the noise matrix, respectively. The instrumental and propagation parameters of stations p and q along the direction of the m -th source are represented by \mathbf{J}_{pm} and \mathbf{J}_{qm} ($\in \mathbb{C}^{2 \times 2}$), respectively. The source coherency matrix of the m -th source is given by \mathbf{C}_{pqm} ($\in \mathbb{C}^{2 \times 2}$). Throughout this paper, we assume all the sources are unpolarized and therefore, \mathbf{C}_{pqm} to be diagonal. Calibration is the estimation of \mathbf{J}_{pm} for all $p \in [1, N]$ and $m \in [1, M]$. As noted by Hamaker (2000), for any unitary matrix \mathbf{U}_m ($\in \mathbb{C}^{2 \times 2}$), $\mathbf{U}_m^H \mathbf{U}_m = \mathbf{U}_m \mathbf{U}_m^H = \mathbf{I}$, a valid calibration solution would also be $\mathbf{J}_{pm} \mathbf{U}_m$. Note that this unitary ambiguity \mathbf{U}_m can have different values for different m , or different directions.

Even with a solution that has an ambiguity, the data could still be corrected for the estimated errors because the ambiguity will cancel out during correction. However, the ambiguity prevents us from using the calibration solutions to model the instrument (such as the beam shape) or propagation phenomena (such as the ones that happen in the ionosphere). The simplest example of further use of solutions is finding their mean. Let us consider having two solutions for station p along the direction m , say at adjacent time or frequency intervals: $\mathbf{J}_{pm1} \mathbf{U}_{m1}$ and $\mathbf{J}_{pm2} \mathbf{U}_{m2}$. Also assume that the true values of the Jones matrices are almost identical, i.e., $\mathbf{J}_{pm1} \approx \mathbf{J}_{pm2} = \mathbf{J}_{pm}$. The true mean is $(\mathbf{J}_{pm1} + \mathbf{J}_{pm2})/2 = \mathbf{J}_{pm}$ while we obtain the interpolant $(\mathbf{J}_{pm1} \mathbf{U}_{m1} + \mathbf{J}_{pm2} \mathbf{U}_{m2})/2 = \mathbf{J}_{pm} (\mathbf{U}_{m1} + \mathbf{U}_{m2})/2$. Due to the fact that $(\mathbf{U}_{m1} + \mathbf{U}_{m2})/2$ is not a unitary matrix ($\mathbf{U}_{m1}^H \mathbf{U}_{m2} \neq \mathbf{0}$), the interpolant is not a valid solution for \mathbf{J}_{pm} anymore and will not satisfy (1).

We can rewrite (1) as

$$\mathbf{V}_{pq} = \sum_{m=1}^M \mathbf{A}_p \mathbf{J}_m \mathbf{C}_{pqm} \mathbf{J}_m^H \mathbf{A}_q^T + \mathbf{N}_{pq} \quad (2)$$

where \mathbf{J}_m is the augmented matrix of all Jones matrices along the m -th direction

$$\mathbf{J}_m \triangleq [\mathbf{J}_{1m}^T, \dots, \mathbf{J}_{Nm}^T]^T, \in \mathbb{C}^{2N \times 2}. \quad (3)$$

The canonical selection matrix \mathbf{A}_p (and \mathbf{A}_q likewise) is given as

$$\mathbf{A}_p \triangleq [\mathbf{0}, \mathbf{0}, \dots, \mathbf{I}, \dots, \mathbf{0}], \in \mathbb{C}^{2 \times 2N}. \quad (4)$$

In (4), the p -th block of columns is a 2×2 identity matrix while the rest is all 0. Then, we see that $\mathbf{J}_m \mathbf{U}_m$ (where \mathbf{U}_m is unitary) is a valid solution for \mathbf{J}_m .

In this paper, we try to solve the following problem: Given a set of solutions whose intrinsic values (i.e., solutions without an ambiguity) are almost equal, we find the interpolant that is immune to unitary ambiguities. Let the set of solutions (taken at adjacent time and frequency intervals and even along adjacent directions) be \mathcal{S} ,

$$\mathcal{S} = \{\mathbf{J}_1, \mathbf{J}_2, \dots, \mathbf{J}_K\} \quad (5)$$

that has K elements. The elements in \mathcal{S} satisfy

$$\mathbf{J}_k = \tilde{\mathbf{J}}_k \mathbf{U}_k, \quad k \in [1, K] \quad (6)$$

where $\tilde{\mathbf{J}}_k$ is the intrinsic value of solution \mathbf{J}_k and \mathbf{U}_k is the unitary ambiguity. We make the additional assumption that all intrinsic values are almost the same, i.e.,

$$\tilde{\mathbf{J}}_1 \approx \tilde{\mathbf{J}}_2 \dots \approx \tilde{\mathbf{J}}_K. \quad (7)$$

This assumption mostly holds for solutions obtained at adjacent time and frequency intervals as well as along adjacent directions, provided that the scale difference due to source models along adjacent directions is taken into account. We make this assumption so that we are not affected by any aliasing errors. Note that the unitary matrices are almost surely not equal

$$\mathbf{U}_1 \neq \mathbf{U}_2 \dots \neq \mathbf{U}_K. \quad (8)$$

We restate the problem we need to solve: Given the set of solutions \mathcal{S} , find the mean $\bar{\mathbf{J}}$ of the solutions such that it is the most accurate approximation of the intrinsic mean

$$\bar{\mathbf{J}} = \frac{1}{K} \sum_{k=1}^K \tilde{\mathbf{J}}_k \quad (9)$$

within a unitary ambiguity. By replacing the summation with weighted summation in (9), this could also be extended to interpolation with any positive weighting scheme.

3 INTERPOLATION

In this section, we first present the well established scalar averaging technique used in radio interferometry. As mentioned before, this does not extend to the case with Jones matrices as calibration solutions. In order to proceed further, we give a rather pedagogical overview of the manifold geometry of calibration solutions. Next, we present an averaging algorithm (Algorithm I) using the quotient manifold in section 3.3. We also give an alternative algorithm (mainly for comparison purposes) that averages using the tangent space of the quotient manifold (Algorithm II) in section 3.4.

3.1 Interpolation: The scalar case

The calibration solutions of N stations for an interferometer with a single polarization can be given as

$$\mathbf{g} = \begin{bmatrix} g_1 \\ g_2 \\ \vdots \\ g_N \end{bmatrix} e^{j\psi} = \begin{bmatrix} |g_1| e^{j(\angle g_1 + \psi)} \\ |g_2| e^{j(\angle g_2 + \psi)} \\ \vdots \\ |g_N| e^{j(\angle g_N + \psi)} \end{bmatrix} \quad (10)$$

where ψ is the phase ambiguity common to all stations. Consider the averaging of K such solutions given by the set \mathcal{G} as

$$\mathcal{G} = \{\mathbf{g}_1 e^{j\psi_1}, \mathbf{g}_2 e^{j\psi_2}, \dots, \mathbf{g}_K e^{j\psi_K}\}. \quad (11)$$

Consider the calculation of the average phase for station n using the solutions in the set \mathcal{G} . Normally, we keep one station (out of N) as the reference (say the first station). With one station kept as the reference, the average phase of station n becomes

$$\begin{aligned} \overline{\angle g_n} &= \frac{1}{K} \sum_{k=1}^K (\angle g_{nk} + \psi_k - \angle g_{1k} - \psi_k) \\ &= \frac{1}{K} \sum_{k=1}^K \angle g_{nk} - \frac{1}{K} \sum_{k=1}^K \angle g_{1k} \end{aligned} \quad (12)$$

where $\frac{1}{K} \sum_{k=1}^K \angle g_{nk}$ is the intrinsic average. Moreover, the term $\frac{1}{K} \sum_{k=1}^K \angle g_{1k}$ is common to the averaged phases of all stations. Therefore, phase averaging can be solved for the single polarized case upto a common phase ambiguity. As the ambiguity $e^{j\psi}$ does not affect the amplitudes of the solutions, amplitudes can also be averaged without hindrance. Furthermore, by using positive weights in the summation of (12), the same method can be applied to any interpolation scheme. This form of averaging and interpolation is widely used in current interferometric data processing. However, this method does not extend to calibration solutions with dual polarized interferometers where we have Jones matrices as our solutions.

3.2 The quotient manifold structure of calibration solutions

We provide a brief overview of manifolds and Lie groups before we proceed. A more general overview of this subject can be found in Tu (2011) and Absil et al. (2008). A manifold can be loosely described as a set of entities, together with a set of mappings (charts) that can locally describe the manifold in Euclidean space. A “quotient” manifold is a sub-manifold of a larger manifold and the entities in the quotient manifold represent more than one entity in the embedding manifold.

This notion of a quotient manifold naturally represents the calibration solutions with unitary ambiguities. Given the set \mathcal{S} in (5), we consider two solutions (say \mathbf{J}_1 and \mathbf{J}_2) to be “similar” if they are related by a unitary matrix, i.e.,

$$\mathbf{J}_1 \sim \mathbf{J}_2 \Leftrightarrow \mathbf{J}_1 = \mathbf{J}_2 \mathbf{U}_{12} \quad (13)$$

where \mathbf{U}_{12} is unitary. The equivalence relation \sim satisfies reflexive, symmetric, and transitive conditions as described in Absil et al. (2008). Therefore, assuming the intrinsic value of all elements in \mathcal{S} are the same (7), we can select only

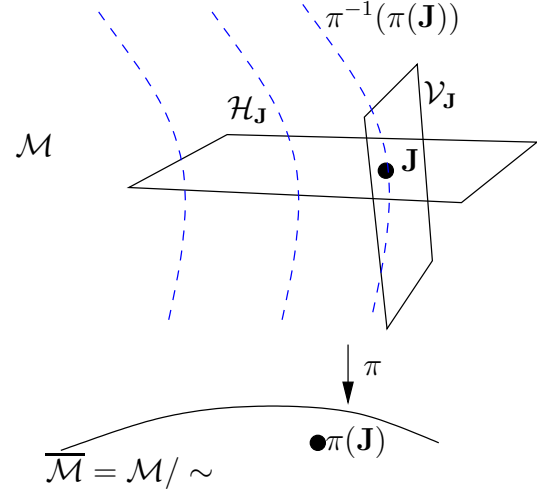


Figure 1. The quotient manifold geometry (Absil et al. 2008) of the calibration solutions. The dashed (blue) line (on \mathcal{M}) represents the equivalence class of all solutions that are related to \mathbf{J} by a unitary ambiguity. This equivalence class is represented by a single point on the quotient manifold $\overline{\mathcal{M}} = \mathcal{M}/\sim$. The vertical space $\mathcal{V}_{\mathbf{J}}$ is the vector space tangential to the equivalence class and the horizontal space $\mathcal{H}_{\mathbf{J}}$ is the orthogonal complement.

one element from \mathcal{S} to represent the whole set, under the equivalence relation \sim , given by (13).

Consider \mathcal{M} to be the manifold of all $2N \times 2$ complex matrices ($\mathbb{C}^{2N \times 2}$), then, we can represent all elements in \mathcal{S} on the quotient manifold $\overline{\mathcal{M}} = \mathcal{M}/\sim$, where the equivalence relation is given by (13). The mapping π (canonical projection) is defined such that any matrix $\mathbf{J}\mathbf{U}$ on \mathcal{M} (\mathbf{U} unitary) is mapped onto a single point, $\pi(\mathbf{J})$ on \mathcal{M}/\sim .

With the mapping π , we have the equivalence class

$$\pi^{-1}(\pi(\mathbf{J})) = \{\mathbf{J}\mathbf{U} : \mathbf{U}\mathbf{U}^H = \mathbf{U}^H\mathbf{U} = \mathbf{I}, \mathbf{U} \in \mathbb{C}^{2 \times 2}\} \quad (14)$$

of solutions that is represented by a single point on \mathcal{M}/\sim , as illustrated in Fig. 1.

The vertical space and horizontal space are vector spaces that are related to the manifold as follows. We take the vertical space to be the tangent space to the equivalence class $\pi^{-1}(\pi(\mathbf{J}))$ at \mathbf{J} , i.e.,

$$\mathcal{V}_{\mathbf{J}} = \{\mathbf{J}\mathbf{A} : \mathbf{A}^H = -\mathbf{A}, \mathbf{A} \in \mathbb{C}^{2 \times 2}\} \quad (15)$$

and we choose the orthogonal complement of the vertical space as the horizontal space $\mathcal{H}_{\mathbf{J}}$,

$$\mathcal{H}_{\mathbf{J}} = \{\mathbf{J}_{\perp} \mathbf{B} : \mathbf{B} \in \mathbb{C}^{2N-2 \times 2}\}. \quad (16)$$

The orthogonal complement \mathbf{J}_{\perp} ($\in \mathbb{C}^{2N \times 2N-2}$) is a matrix whose columns are orthogonal to those of the matrix \mathbf{J} , i.e., $\mathbf{J}^H \mathbf{J}_{\perp} = \mathbf{0}$.

With this formal representation of the manifold geometry of the solution space at hand, we present two algorithms for interpolation in sections 3.3 and 3.4. Similar averaging techniques on other manifolds are already being investigated. For instance, Kaneko et al. (2012) present an algorithm for averaging on a compact Stiefel manifold. A similar algorithm for averaging on the Grassmann manifold is given by Amsellem & Farhat (2008). While a manifold has a more geometric structure, a Lie group has a more algebraic structure. A Lie group can be described as a set of entities, with an identity element and operations for multiplication and

inverse (Tu 2011). There is a close relation between smooth manifolds and Lie groups and in Fiori & Tanaka (2009), several averaging techniques for square matrices using the Lie group structure is presented. In Fiori (2011), the Lie group structure as well as the manifold geometry is exploited for averaging.

The algorithm presented in section 3.3 (Algorithm I), performs the averaging directly on the quotient manifold while the algorithm presented in section 3.4 (Algorithm II), performs the averaging in the tangent space to the manifold. Algorithm II is presented mainly for comparison with Algorithm I, as most existing work on averaging, such as (Kaneko et al. 2012; Amsellem & Farhat 2008), is done in the tangent space. Both these algorithms can be extended to interpolation by using any positive weighting scheme.

3.3 Averaging on the quotient manifold (Algorithm I)

We present the algorithm to find the mean of the set \mathcal{S} in (5). Let us call the estimated mean as $\bar{\mathbf{J}}$. The basic idea is to find the element on the quotient manifold that represents the set \mathcal{S} as accurately as possible. This would also be the mean of the set \mathcal{S} .

- (i) Start with the initial estimate as one value from \mathcal{S} , as $\bar{\mathbf{J}} \leftarrow \mathbf{J}_1$. The error threshold is given by ϵ .
- (ii) For each element in \mathcal{S} , find unitary $\mathbf{P}_k \in (\mathbb{C}^{2 \times 2})$ such that

$$\mathbf{P}_k = \arg \min_{\mathbf{P}_k, \mathbf{P}_k^H \mathbf{P}_k = \mathbf{P}_k \mathbf{P}_k^H = \mathbf{I}} \|\bar{\mathbf{J}} - \mathbf{J}_k \mathbf{P}_k\|^2. \quad (17)$$

This is basically an “alignment” operation and the details of this step are given in section 3.5.

- (iii) Form

$$\bar{\mathbf{G}} = \frac{1}{K} \sum_{k=1}^K \mathbf{J}_k \mathbf{P}_k. \quad (18)$$

- (iv) Find the unitary projector \mathbf{P} to minimize $\|\bar{\mathbf{J}} - \bar{\mathbf{G}}\mathbf{P}\|^2$ as given in section 3.5.
- (v) If $\|\bar{\mathbf{J}} - \bar{\mathbf{G}}\mathbf{P}\| < \epsilon$ then stop. Else update $\bar{\mathbf{J}} \leftarrow \bar{\mathbf{G}}$ and go to step (ii).
- (vi) Return $\bar{\mathbf{J}}$ as the mean.

Note that this algorithm can be modified for interpolation by replacing the averages in (18) with weighted averaging using positive weights. The proof of convergence of this algorithm is part of future research. For the moment, we rely on numerical simulations to test its convergence in section 4.

3.4 Averaging in the tangent space (Algorithm II)

The algorithm presented in this section projects each element in the set \mathcal{S} to the horizontal space of the quotient manifold before averaging is performed. As before, let us call the estimated mean as $\bar{\mathbf{J}}$.

- (i) Start with the initial estimate as one value from \mathcal{S} , as $\bar{\mathbf{J}} \leftarrow \mathbf{J}_1$. The error threshold is given by ϵ . For better convergence, a positive scalar $\rho \in (0, 1]$ is used.

- (ii) Form the orthogonal projector matrix $\mathbf{V} \in \mathbb{C}^{2N \times 2N}$

$$\mathbf{V} = \mathbf{I} - \bar{\mathbf{J}}(\bar{\mathbf{J}}^H \bar{\mathbf{J}})^{-1} \bar{\mathbf{J}}^H. \quad (19)$$

- (iii) Project each element in \mathcal{S} , onto the horizontal space $\mathcal{H}_{\bar{\mathbf{J}}}$ as

$$\mathbf{W}_k = \mathbf{V} \mathbf{J}_k. \quad (20)$$

- (iv) Form the average in $\mathcal{H}_{\bar{\mathbf{J}}}$ as

$$\bar{\mathbf{W}} = \frac{1}{K} \sum_{k=1}^K \mathbf{W}_k. \quad (21)$$

- (v) Form current estimate for the average as $\bar{\mathbf{G}} = \bar{\mathbf{J}} + \rho \bar{\mathbf{W}}$. Find the unitary projector \mathbf{P} to minimize $\|\bar{\mathbf{J}} - \bar{\mathbf{G}}\mathbf{P}\|^2$ as given in section 3.5.

- (vi) If $\|\bar{\mathbf{J}} - \bar{\mathbf{G}}\mathbf{P}\| < \epsilon$ then stop. Else update $\bar{\mathbf{J}} \leftarrow \bar{\mathbf{G}}$ and go to step (ii).

- (vii) Return $\bar{\mathbf{J}}$ as the mean.

3.5 Finding unitary projector to minimize $\|\bar{\mathbf{J}} - \mathbf{J}_k \mathbf{P}_k\|^2$

What we have to solve is in fact the matrix Procrustes problem (Schönemann 1966) and we use the algorithm given in Higham (2008).

- (i) Find the product $\mathbf{X} = \mathbf{J}_k^H \bar{\mathbf{J}}$.
- (ii) Find the singular value decomposition of \mathbf{X} as

$$\mathbf{U} \mathbf{S} \mathbf{V}^H = \mathbf{X}. \quad (22)$$

- (iii) Return $\mathbf{P}_k = \mathbf{U} \mathbf{V}^H$.

The proof can be found in Higham (2008).

3.6 Discussion

The main assumption used in Algorithm I is that all elements in \mathcal{S} belong to the equivalence class (dashed line in Fig. 1) or are very close to it. So on the quotient manifold, they lie on a small area that can be considered locally Euclidean, thus enabling averaging. For Algorithm II, we perform the averaging in the tangent space, which is a vector (Euclidean) space and therefore averaging works. Computationally, Algorithm II is much more expensive because of the calculation of the projection matrix at each iteration. Also, it is numerically less stable and hence the need of ρ for better convergence. We use Algorithm II mainly for comparing the performance of Algorithm I. The fundamental question posed here is whether (for our specific problem) it is better to average on the quotient manifold or in the tangent space. Most existing work on averaging or interpolation use operations in the tangent space therefore comparison of Algorithms I and II in terms of their performance is important. In the next section, we give simulation results to compare their performance.

4 SIMULATION RESULTS

We consider an interferometric observation with $N = 40$ stations and interpolation of $K = 10$ solutions. Therefore the set \mathcal{S} in (5) has cardinality of 10, with each matrix \mathbf{J}_k of size 80×2 . The intrinsic values of the first matrix, $\bar{\mathbf{J}}_1$,

are generated as $\mathcal{U}(0, 1) + j\mathcal{U}(0, 1)$ (drawn from a uniform distribution in $[0, 1]$). The intrinsic values of the remaining matrices $\tilde{\mathbf{J}}_2, \dots, \tilde{\mathbf{J}}_K$ are obtained by perturbing the elements of $\tilde{\mathbf{J}}_1$ by adding $\sigma(\mathcal{U}(0, 1) + j\mathcal{U}(0, 1))$ to them. The value for σ is varied for different simulations as explained later. However, we keep the value of σ in $[0, 0.5]$ to ensure the intrinsic variation is small enough that averaging (or interpolation) is not dominated by aliasing. Once the matrices are generated in this fashion, we multiply each intrinsic matrix $\tilde{\mathbf{J}}_k$ by a random unitary matrix \mathbf{U}_k to get $\mathbf{J}_k = \tilde{\mathbf{J}}_k \mathbf{U}_k$. To each realization of \mathbf{J}_k , a noise matrix \mathbf{N} is added. The elements of \mathbf{N} are generated to be complex circular Gaussian random variables. The values of \mathbf{N} are scaled to get a given signal to noise ratio (SNR) where the SNR is defined by

$$\text{SNR} \triangleq \frac{\|\mathbf{J}_k\|^2}{\|\mathbf{N}\|^2} \quad (23)$$

before adding them to \mathbf{J}_k .

We also calculate the normal (or Euclidean) average for comparison, as

$$\hat{\mathbf{J}} = \frac{1}{K} \sum_{k=1}^K \mathbf{J}_k. \quad (24)$$

Moreover, we also calculate the intrinsic sample variance, using the intrinsic mean of (9) as

$$\text{var}(\bar{\mathbf{J}}) = \frac{1}{K\|\bar{\mathbf{J}}\|^2} \sum_{k=1}^K \|\bar{\mathbf{J}} - \tilde{\mathbf{J}}_k\|^2. \quad (25)$$

The criterion that we use for measuring the performance of the various averaging algorithms is the normalized error (NE), defined as

$$\text{NE} \triangleq \frac{\|\bar{\mathbf{J}} - \bar{\mathbf{J}}\mathbf{U}\|^2}{\|\bar{\mathbf{J}}\|^2} \quad (26)$$

where $\bar{\mathbf{J}}$ is the intrinsic mean, $\bar{\mathbf{J}}$ is the estimated mean, using (i) Algorithm I (ii) Algorithm II or, (iii) Euclidean average of (24) and finally, \mathbf{U} is a unitary projector determined as in section 3.5 to minimize $\|\bar{\mathbf{J}} - \bar{\mathbf{J}}\mathbf{U}\|^2$.

4.1 Simulation I

We generate the set \mathcal{S} as described above 100 times, keeping $\sigma = 0.1$ and $\text{SNR} = 100$. For each realization, we estimate the average by Algorithm I, using 10 iterations and Algorithm II, using 60 iterations (the reason for this will be explained later) with $\rho = 0.3$. We have shown the normalized estimation error (NE) in Fig. 2 for different approaches. We have also plotted the intrinsic variance, calculated using (25) in this figure. As seen in this figure, the proposed Algorithm I has an error almost comparable with the intrinsic variance. The proposed Algorithm II has a slightly worse performance but it is still better than Euclidean averaging.

In Fig. 3 and Fig. 4, we have shown the convergence performance of both Algorithm I and Algorithm II. We measure the convergence by the norm of the difference between the current estimate and the updated estimate at each iteration. It is clear that Algorithm I has much better convergence (only about 4 iterations) than Algorithm II, which does not converge even after 60 iterations.

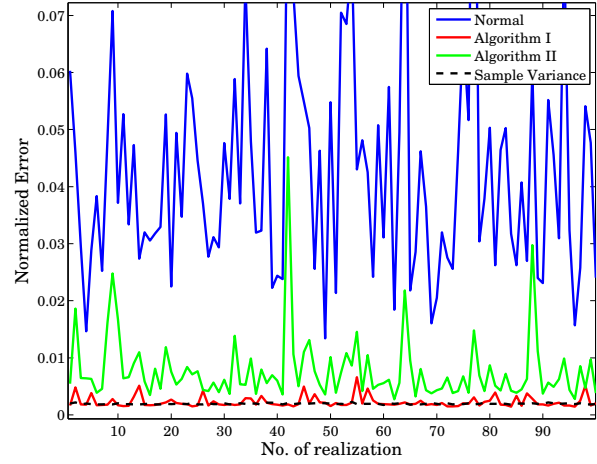


Figure 2. Normalized error for 100 realizations of \mathcal{S} with $\sigma = 0.1$ and $\text{SNR} = 100$. The intrinsic sample variance is given by the dashed line. The proposed Algorithm I performs much better than normal (Euclidean) averaging and almost at the level of intrinsic variance. Algorithm II performs better than Euclidean averaging but is worse than Algorithm I.

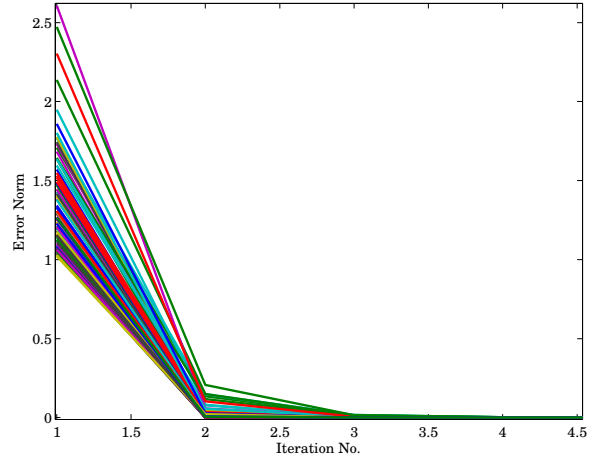


Figure 3. The convergence of Algorithm I for 100 different realizations. After about 4 iterations, no further improvement occurs and therefore, we can fix the maximum number of iterations at 4.

4.2 Simulation II

We vary the values of σ and SNR in this simulation. For each value of σ and SNR, we generate 100 realizations of the set \mathcal{S} as before and compute the average. In Fig. 5, we have shown the mean normalized error over all realizations for both proposed algorithms as well as for Euclidean averaging. For Algorithm I, we used 4 iterations and for Algorithm II, we used 60 iterations with $\rho = 0.3$. From Fig. 5 it is clear that Algorithm I gives the best results. Moreover, the dependence on σ is not significant. The performance of all three methods degrade at low values of SNR. Algorithm II performs worse than Algorithm I and comparing the computational cost, Algorithm I clearly gives better results.

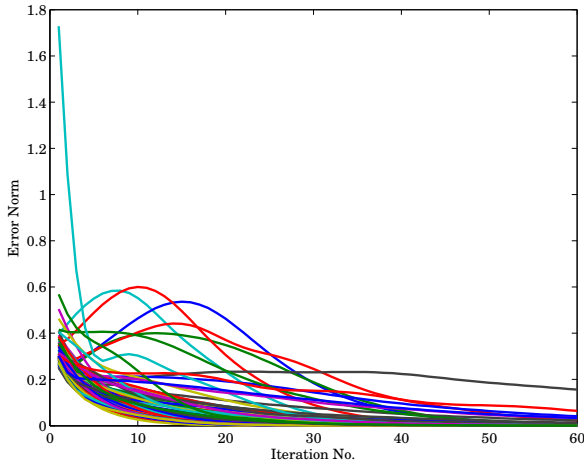


Figure 4. The convergence of Algorithm II for 100 different realizations. Even after 60 iterations there are some cases that do not converge.

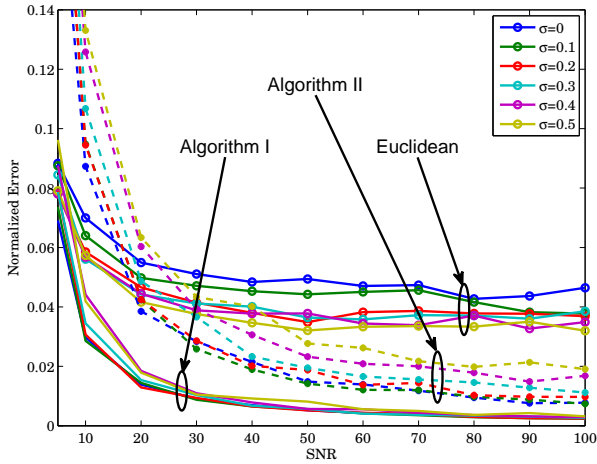


Figure 5. The mean normalized error over 100 realizations for various values of σ and SNR. Algorithm I performs significantly better than normal (Euclidean) averaging while Algorithm II performs slightly better but worse than Algorithm I. Algorithm I is almost insensitive to the values of σ , especially at high values of SNR.

4.3 Simulation III

In this simulation, we perform weighted averaging, where the weights are generated from a uniform distribution in $[0, 1]$. Once again, for each value of σ and SNR, we generate 100 realizations of \mathcal{S} and for each realization, we generate a new set of weights. We emphasize that Algorithm II did not give convergent results even at 60 iterations and for any value of ρ . Therefore, we omitted the results of Algorithm II because it clearly fails to perform weighted averaging.

5 CONCLUSIONS

We have presented a method for averaging Jones matrices obtained in radio interferometric calibration by exploit-

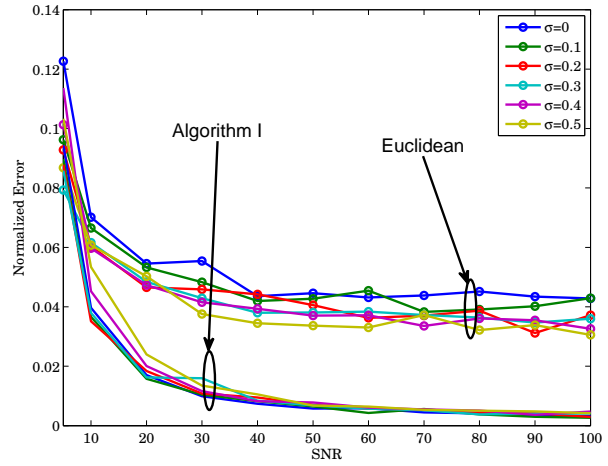


Figure 6. The mean normalized error over 100 realizations for various values of σ and SNR under weighted averaging. Algorithm I performs better than Euclidean averaging while Algorithm II fails to perform well.

ing the quotient manifold structure of such solutions. This method could also be used for weighted averaging, or interpolation. Unlike Euclidean averaging, which gives inaccurate results due to the unknown unitary ambiguity in the solutions, the proposed methods give significantly better results. For comparison, we have also proposed an alternative algorithm that operates in the tangent space to the manifold. Simulation results suggest that averaging directly on the quotient manifold is better than averaging by using the tangent space, for our specific problem. One possible reason for this could be due to numerical instability. Existing work that exploit the tangent space for averaging have matrices with orthonormal columns (e.g. the Stiefel manifold). However, in our problem we do not have such a constraint and this could result in numerical instability. Future work would focus on improving the numerical stability and interpolating along geodesics of the quotient manifold.

ACKNOWLEDGMENTS

We thank the reviewers: Simone Fiori and Yves Wiaux for a careful review and valuable comments that helped to enhance this paper.

REFERENCES

- Absil P. A., Mahony R., Sepulchre R., 2008, Optimization Algorithms on Matrix Manifolds. Princeton Univ. Press, Princeton NJ
- Amsellem D., Farhat C., 2008, AIAA Journal, 46, 1803
- Fiori S., 2011, Front. Electr. Eng. China, 6, no. 1, 137
- Fiori S., Tanaka T., 2009, IEEE Trans. on Sig. Proc., 57, no. 12, 4734
- Hamaker J. P., 2000, Astronomy and Astrophysics Supp., 143, 515
- Hamaker J. P., Bregman J. D., Sault R. J., 1996, Astronomy and Astrophysics Supp., 117, 96

- Higham N., 2008, Functions of matrices: Theory and computation. Philadelphia USA: SIAM
- Intema H. T., van der Tol S., Cotton W. D., Cohen A. S., van Bemmelen I. M., Rottgering H. J. A., 2009, A&A, 501, 1185
- Kaneko T., Tanaka T., Fiori S., 2012, in proc. IEEE ICASSP, Kyoto, Japan, pp 3829–3832
- Kazemi S., Yatawatta S., Zaroubi S., Labropoulos P., de Bruyn G., Koopmans L., Noordam J., 2011, MNRAS, 414, no. 2, 1656
- Schönemann P., 1966, Psychometrika, 31, no.1, 1
- Tu L., 2011, An Introduction to Manifolds. Springer
- Yatawatta S., 2012, in proc. IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM), Hoboken NJ, pp 533–536